A SYSTEM AND METHOD FOR MULTI-MODAL AUTHENTICATION
USING SPEAKER VERIFICATION

Background of the Invention

5   ### Field of the Invention

The present invention relates generally to electronic transactions, and particularly to verifying and authenticating electronic transactions.

Technical Background

10   The use and uses of electronic transactions in commerce are ubiquitous. Many transactions are being conducted on-line, between users and commercial web-sites operating in the electronic market place. These web-sites are sponsored by banks, stock brokerage firms, retailers, wholesalers and countless others. Other transactions are being conducted using point-of-sale (POS) terminals in brick and mortar commercial establishments. Often

15   POS terminals include credit, debit, and check authorization capabilities. Still other transactions, such as cash withdrawals, are being conducted using ATM machines provided by financial institutions. Some of these devices are used as stand-alone devices and some are networked. Because of the sheer magnitude of cash being transferred electronically, security is absolutely critical. Both financial and commercial institutions are concerned with the

20   difficulty in obtaining verification and authentication during such transactions. Stolen credit cards are often used by criminal elements to fraudulently purchase goods and services, withdraw cash, or conduct other financial transactions. Computer hackers are also a threat.

What is needed is a secure system and method for authenticating and verifying the identity of the parties involved in an electronic transaction. What is needed is a system and method for substantially eliminating the fraudulent usage of debit and credit cards during electronic transactions. A method and system for authentication is needed to provide security

during on-line transactions, ATM transactions, and point-of-sale (POS) transactions.

Summary of the Invention

The present invention addresses the needs described above. The present invention provides a secure system and method for providing authenticating and verifying data during the course of an electronic transaction.

One aspect of the present invention is a computerized method for authenticating an electronic transaction between a user and a computer. The computer is configured to conduct electronic transactions. The method includes: receiving a computer-generated transaction identifier from the computer via an electronic data link; receiving a user-spoken transaction identifier and a user-spoken verification identifier transmitted by the user via a voice connection; comparing the user-spoken transaction identifier with the computer transaction identifier; comparing the user-spoken verification identifier with a voice print of the user; and transmitting an authentication message to the computer if the user-spoken transaction identifier matches the computer-generated transaction identifier and if the user-spoken verification identifier matches the voice print.

In another aspect, the present invention includes a system for authenticating an electronic transaction between a first user-operated device and a computer. The computer is configured to conduct electronic transactions. The system includes a voice browser configured to receive and process user-spoken information when coupled to a second user-operated device. The voice browser is programmed to compare a user-spoken transaction identifier to a computer-generated transaction identifier, and to compare a user-spoken verification identifier to a voice print of the user. A session correlator is coupled

2

to the voice browser. The session correlator is configured to transmit an authentication message to the computer if the user-spoken transaction identifier matches the computer transaction identifier, and if the user-spoken verification identifier matches the voice print.

In another aspect, the present invention includes a computerized voice verification method for authenticating an electronic transaction between a user and a computer. The computer is configured to conduct electronic transactions. The method includes: enrolling the user in a voice verification system, whereby the user provides the system with a user voice print; performing the electronic transaction; receiving a transaction identifier from the computer via an electronic data link in response to performing the electronic transaction; receiving a user-spoken transaction identifier and a user-spoken verification identifier transmitted by the user via a voice connection; comparing the user-spoken transaction identifier with the computer transaction identifier and the user-spoken verification identifier with a voice print of the user; and transmitting an authentication message to the computer if the user-spoken transaction identifier matches the computer transaction identifier, and if the user-spoken verification identifier matches the voice print.

In yet another aspect, the present invention includes a computerized method for controlling web-site navigation. The method includes: providing an authentication system including a voice recognition unit and a session correlator, the voice recognition unit having access to a pre-registered voice print of the user, whereby the authentication system is coupled to a user computer and a web-site during the computerized method; conducting a transaction between the user computer and the web-site, the web-site transmitting a transaction identifier to the user computer and the authentication system in response to the transaction; receiving a user-spoken transaction identifier and a user-spoken verification

3

identifier via a telephonic connection, the authentication system being programmed to compare the user-spoken transaction identifier to the transaction identifier and the user-spoken verification identifier to the pre-registered voice print; transmitting an authentication message to the web-site if the user-spoken transaction identifier matches the

5   transaction identifier and if the user-spoken verification identifier matches the voice print; receiving at least one user-spoken command for controlling web-site navigation, the authentication system being programmed to convert the at least one user-spoken command into at least one computer-readable command; and transmitting the at least one computer readable command to the web-site, the at least one computer readable command being

10   executed by the web-site, whereby the user controls web-site navigation of the web-site by the at least one user-spoken command.

Additional features and advantages of the invention will be set forth in the detailed description which follows, and in part will be readily apparent to those skilled in the art from that description or recognized by practicing the invention as described herein, including the

15   detailed description which follows, the claims, as well as the appended drawings.

It is to be understood that both the foregoing general description and the following detailed description are merely exemplary of the invention, and are intended to provide an overview or framework for understanding the nature and character of the invention as it is claimed. The accompanying drawings are included to provide a further understanding of the

20   invention, and are incorporated in and constitute a part of this specification. The drawings illustrate various embodiments of the invention, and together with the description serve to explain the principles and operation of the invention.

4

Brief Description of the Drawings

Figure 1 is a block diagram of the authentication system in accordance with a first embodiment of the present invention;

Figure 2 is a block diagram of the authentication system being used in an Internet web-site application in accordance with a second embodiment of the present invention;

Figure 3 is a block diagram of the authentication system being used in an ATM banking application in accordance with a third embodiment of the present invention;

Figure 4 is a block diagram of the authentication system being used in a point-of-sale application in accordance with a fourth embodiment of the present invention;

Figure 5 is a flow chart of an authentication process in accordance with the present invention;

Figure 6 is a flow chart showing a method for controlling a web-site using telephonic voice menu commands; and

Figure 7 is a flow chart showing a method for controlling a web-site using a predetermined suite of voice commands.

Detailed Description of the Invention

Reference will now be made in detail to the present exemplary embodiments of the invention, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the drawings to refer to the same or like parts. An exemplary embodiment of the authentication system of the present invention is shown in Figure 1, and is designated generally throughout by reference numeral 10.

In accordance with the invention, the present invention includes a method and system for authenticating an electronic transaction between a first user-operated device, such as a personal computer, and a commercial institution computer, such as a web-site, configured to conduct electronic transactions. The system includes a voice browser coupled to a second

5   user-operated device, such as a telephone set. The voice browser is configured to receive and process user-spoken information from the second user operated device, whereby a user-spoken transaction identifier is compared to a transaction identifier, and a user-spoken verification identifier is compared to a voice print of the user. A session correlator is coupled to the voice browser. The session correlator is configured to transmit an

10   authentication message to the computer if the user-spoken transaction identifier matches the computer transaction identifier and if the user-spoken verification identifier matches the voice print.

The system and method of the present invention provides secure authentication and verification of user provided data during the course of an electronic transaction. The system and method of the present invention substantially eliminates the fraudulent usage of debit

15   and credit cards during electronic transactions. The system and method of the present invention is effective in providing security during on-line transactions, ATM transactions, and point-of-sale (POS) transactions. The system and method of the present invention also provides the user with a "hands-free" way of navigating the web using a full-duplex voice communications medium (wire line telephone, wireless telephone, radio, and etc.).

20   As embodied herein, and depicted in Figure 1, a block diagram of the authentication system in accordance with a first embodiment of the present invention is disclosed. Authentication system 10 includes voice browser 20 and session correlator 30. Voice

6

browser 20 includes telephony interface 200 connected to telephone network 12 and computer 202. Computer 202 is coupled to voice print database 204, voice menu option library 206, speech synthesizer 208, and speech recognition unit 210. Session correlator 30 includes computer 300, server 302, and session correlator software 304. In the example shown in Figure 1, voice browser 20 and session correlator 30 are network resources disposed anywhere in the network backbone. One of ordinary skill in the pertinent art will recognize that browser 20 and correlator 30 may be co-located in a network data center. In that embodiment, computer 202 and computer 300 may well be embodied in one computer.

It will be apparent to those of ordinary skill in the pertinent art that modifications and variations can be made to telephone interface 200 of the present invention depending on network 12. For example, if interface 200 is connected to a T-1 line, interface 200 must accommodate a bandwidth of about 1.5 Mb/s and 24 64 kb/s voice grade channels. In another embodiment, interface 200 is connected to several T-1 lines. In yet another embodiment, interface 200 is connected to a T-3 line. In this embodiment, interface 200 must accommodate a bandwidth of approximately 45 mb/s and about 672 64 kb/s voice grade channels. In another embodiment, the telephone network is a wireless network, in which case interface 200 must be configured to transmit and receive RF signals, and programmed to accommodate wireless access protocol (WAP). In another embodiment, the telephone network is an I/P network and interface 200 must accommodate a voice-over-I/P protocol such the session initiation protocol (SIP).

It will be apparent to those of ordinary skill in the pertinent art that modifications and variations can be made to server 302 depending on system component choices. One of ordinary skill in the art will recognize that Internet 14 includes physical devices such as

wires, cables, optical fiber, photonic components, routers, bridges, intranets, extranets, and other networks. Server 302 must be configured accordingly. Internet 14 also represents a communications medium that supports standard web protocols such as HTTP and a secure transport protocol.

5   It will be apparent to those of ordinary skill in the pertinent art that modifications and variations can be made to voice menu option library 206, speech synthesizer 208, speech recognition unit 210, and session correlator software 304 depending on the implementation software used in developing each of these modules. In one embodiment, all of these modules are resident in a Java application server located in a network data center. As

10   discussed above, in other embodiments, session correlator software 304 is located in a separate application server in the network backbone. Voice menu option library 206, speech synthesizer 208, and recognition module 210 may be developed using any suitable scripting software development tool, such as Voice XML, IBM's direct talk, or by using the software tools marketed by Nuance, Inc.

15   It will be apparent to those of ordinary skill in the pertinent art that modifications and variations can be made to session correlator software 304 depending on the degree of sophistication of voice browser 20. In one embodiment, library 206 includes a suite of menu options for providing authentication and verification. In this embodiment, session correlator module 304 is programmed to provide a simple interface between the web-site server and

20   voice browser 20. After a transaction is requested, the web-site server provides the authentication identifier and credit card data to system 10, via session correlator 30. Session correlator 30 provides the web-site server with an authentication message or a denial message, depending on the outcome of the comparison made by voice browser 20. In

8

another embodiment, voice browser 20 is used to navigate web-pages. Thus, the complexity of session correlator module 304 is increased to provide an interface between the commercial computer and voice browser 20. Session correlator module 304 is programmed to provide current web-page data to voice browser 20 in order for voice browser 20 to

5   provide the user with a suite of voice commands that are correlated with icons displayed on the current web-page. Session correlator module 304 is also programmed to transmit each command in the suite of voice commands in a format recognized by the web-site.

One of ordinary skill in the art will also recognize that modifications and variations can be made to voice menu option library 206 depending on the flexibility inherent in the user

10   interface of voice browser 20. In one embodiment, voice menu option library 206 consists of a database of menu options that is used in conjunction with voice print database 204, speech synthesizer 208, and speech recognition software 210 during the authentication process. Browser computer 202 accesses library 206 to obtain the appropriate user prompt. Subsequently, the user is prompted for the proper transaction identifier and verification

15   identifier. In another embodiment, library 206 is more complex. It includes menu options for controlling web-site navigation by voice command. Speech recognition module 210 is programmed to interpret possible user responses to the synthesized voice menu options. In one embodiment, the menu options are designed to prompt the user to make a selection by using pre-selected words or phrases as suggested by the prompt. In another embodiment, the

20   user employs numeric answers to make menu selections. He speaks "one" when selecting menu option one, "two" when selecting menu option two, and so on.

In yet another embodiment, library 206 includes a suite of navigation commands that allow the user to control web-site navigation. In one version, the suite is designed as a static

9

set of commands. In another version, the software is programmed to provide a dynamic

suite of commands that adapt to changing web-site environments. In both versions, session

correlator 30 provides library 206 with information regarding the web-page that is currently

being accessed by the user. In the static version, computer 202 accesses the pre-loaded static

5   commands in library 206. The static commands relate to cursor movement and mouse

clicking operations. In the dynamic version, computer 202 uses the web-page information to

generate and load a set of commands that reflect the contents of the current web-page. The

dynamic commands allow the user to select any icon by speaking the name of the icon. For

example, if the user is navigating a search-engine, she says "Finance" to select the "Finance"

10   icon displayed on the web-page.

As embodied herein, and depicted in Figure 2, a block diagram of authentication system 10

being used in an Internet web-site application in accordance with a second embodiment of

the present invention is disclosed. In this application, user 40 employs his personal

computer 44 to make an electronic transaction with web-site 50. Personal computer 44 is

15   coupled to web-site 50 via the Internet 14. In this embodiment, the present invention is

used, for example, in purchasing airline tickets, performing on-line banking, or participating

in an on-line auction. Another important application involves music distribution. The

present invention is used to authenticate users so that they may download music files to their

personal computer or their music player. The above applications are representative

20   examples, and the present invention should not be construed as being limited by them.

    In response to the user's request for a transaction, web-site 50 provides the user with

an authentication identifier. The user dials a predetermined number corresponding to the

authentication service to connect his telephone set 42 to voice browser 20 via telephone

10

network 12. When the connection is made, voice browser 20 initiates the call with a voice prompt. In response, user 40 provides voice browser 20 with the transaction identifier received from the web-site, and the pre-registered verification identifier. After providing the user with a transaction identifier, server 52 transmits the transaction identifier to

5   authentication system 10 via the Internet 14. If the two transaction identifiers match, and the verification data provided by the user is correct, session correlator 30 transmits an authentication message to web-site 50 authorizing the transaction. The method of authenticating and verifying an electronic transaction is described in more detail below, in conjunction with the flow diagram shown in Figure 5.

10   As embodied herein, and depicted in Figure 3, a block diagram of the authentication system being used in an ATM banking application in accordance with a third embodiment of the present invention is disclosed. In this embodiment, the user employs ATM machine 60 to perform a financial electronic transaction, such as cash withdrawal or a transfer of funds between bank accounts. ATM machine 60 includes display 62, keypad 64, speaker 66, and

15   microphone 68. ATM machine 60 also includes a slot for debit/credit card insertion, and a cash withdrawal slot. ATM machine is coupled to financial institution 70 by way of data link 16. Speaker 66 and microphone 68 are coupled to voice browser 20 via telephone network 12. Server 72 is coupled to session correlator 30 via the Internet 14. In this embodiment, financial institution 70 employs authentication system 10 to authenticate and

20   verify the transaction in a manner described in detail below. In another embodiment, ATM machine 60 is replaced by a teller operated computer. When a customer requests a transaction at the teller's window, the transaction is conducted between the customer and the teller. To perform the authentication procedure, the customer provides his authentication

11

and verification data by speaking them into a microphone connected to the teller's computer.

One of ordinary skill in the art will recognize that modifications and variations can be made to data link 16 depending on the disposition of ATM machine 60. For example, if ATM machine 60 is located in the lobby of a financial institution, data link 16 may include a direct connection to bank computer 74. If ATM machine 60 is at a remote location, ATM machine 60 may include a modem, in which case data link 16 is connected to server 72 via a telephone network or the Internet.

As embodied herein, and depicted in Figure 4, a block diagram of the authentication system being used in a point-of-sale (POS) application in accordance with a fourth embodiment of the present invention is disclosed. This embodiment is very similar to the ATM embodiment depicted in Figure 3. POS terminal 80 replaces the ATM machine. POS terminal 80 includes microphone 82, speaker 84, credit card reader 86, an optional signature verification area 88, stylus 90, and display 92. In this embodiment, commercial institution 70 employs authentication system 10 to authenticate and verify the purchase in a manner described in detail below.

As embodied herein, and depicted in Figure 5, a flow chart of the authentication process in accordance with the present invention is disclosed. In step s500, the user registers with authentication system 10. Registration includes providing system 10 with the user's name and a user verification identifier, e.g. a voice print that biometrically identifies the user. The voice print is provided by having the user speak a numerical identifier, such as his telephone number during a registration session. The user's name or social security number could also be used when creating the voice print. Computer 202 (Figure 1) uses recognition

12

module 210 to capture the voice print and the other user data, and creates a user file in database 204. In another embodiment, the user also provides system 10 with payment information, which includes a credit/debit card number, and the card's expiration date. In other embodiments more detailed user information is provided, such as user address, e-mail address and other information. In step s502, the user performs the electronic transaction with the commercial computer in accordance with any one of the embodiments depicted in Figures 1-4. After the transaction is conducted, the commercial computer generates a transaction identifier and transmits it to the user. The user must call and authenticate within a predetermined time period or the process flow is directed to step s516, and the transaction is denied. In the embodiment depicted in Figure 5, the time period is five minutes. To authenticate, the user provides voice browser 20 a spoken transaction identifier and a spoken verification identifier via the telephone connection. The spoken transaction identifier must match the transaction identifier generated by the commercial institution's computer or the transaction will not be authorized. Even if the transaction identifiers match, the spoken verification identifier must also match the voice print registered by the user in step s500. If both of the identifiers match system 10 generates an authentication message to the commercial institution in step s514. Upon receipt of the authentication message, the electronic transaction is completed.

As embodied herein, and depicted in Figure 6, a flow chart showing a method for controlling a web-site using telephonic voice menu commands is disclosed. In step s600, the authentication procedure described above is performed. In step s602, voice browser 20 transmits a voice menu to the user via the telephonic connection. In this embodiment, voice browser 20 provides the user with several distinct choices that can be selected by speaking

13

the selected menu option, as described above. After the user makes his selection in step s604, speech recognition module 210 converts the user response into a computer-readable command. In step s608, session correlator 30 transmits the command and a session identifier to the web-site. In one embodiment, as shown in Figure 6, the session identifier is the authentication message identifier itself. This process continues until the user is finished navigating the voice menu.

As embodied herein, and depicted in Figure 7, a flow chart showing a method for controlling a web-site using a predetermined suite of voice commands is disclosed. The chart depicted in Figure 7 assumes that the user is a registered member of system 10. In step s700, the user is provided with a suite of commands that are recognized by system 10. In one embodiment, this step is performed using the voice menu system described above. This information can also be provided on-line, or by using written instructions mailed to the user's home after registration. Before navigating a web-site using verbal commands, the user must perform the authentication described above. In step s704, the user employs a recognized verbal command to perform the desired web-site navigation action. As described above, in one embodiment the user speaks the name of the icon he desires to select. In step s706, speech recognition module 210 converts the spoken command into a machine readable command. Subsequently, correlator 30 transmits the machine-readable command to the web-site for execution. To end the navigation session, the user merely says "end" or some other word indicating that the voice browsing session is over.

It will be apparent to those skilled in the art that various modifications and variations can be made to the present invention without departing from the spirit and scope of the invention. Thus, it is intended that the present invention cover the modifications and

14

variations of this invention provided they come within the scope of the appended claims and their equivalents.